ProNoBiS
Probability and Nondeterminism, Bisimulations and Security

The goals of this proposal fit in the INRIA priority challenge
"**Guaranteeing the reliability and security of software-intensive systems**".

Consortium[1]

Institute	Team	Members		
INRIA Futurs	SECSI	Jean Goubault-Larrecq	(Prof.)	PI and leader
INRIA Futurs	Comète	Catuscia Palamidessi	(DR)	PI
		Peng Wu	(Postdoc)	
		Kostas Chatzikokolakis	(PhD student)	
		Romain Beauxis	(PhD student)	
LSV,	axe Tempo	Laurent Fribourg	(DR)	
		Claudine Picaronny	(MdC)	
		Simon Pinot	(PhD student)	
PPS, U. Paris 7		Vincent Danos	(DR)	PI
		Russell Harmer (MdC)		
U. Birmingham		Marta Kwiatkowska	(Prof.)	PI
U. Verona		Roberto Segala	(Prof.)	PI
		Augusto Parma	(PhD student)	
		Andrea Turrini	(Research Intern)	

[1] There is an isomorphism between the SECSI team at INRIA Futurs and the "axe SECSI" at LSV, ENS Cachan. The separation between INRIA/SECSI and the other teams involved at LSV is therefore essentially an administrative fiction: the three "axes" of LSV, SECSI, Infini, and Tempo should be taken collectively as showing the implication of LSV as a whole in this proposal.

Random selection is too important to be left to chance
Robert R. Coveyou

Introduction

Non-deterministic and probabilistic choice. Most current models of security protocols, e.g., the spi-calculus [2], the applied pi-calculus [1], the seal calculus [6], or the various versions of the Dolev-Yao model [20], are *non-deterministic*: at any state of the protocol, one of several actions can be taken. Typically, honest participants act deterministically, but the attacker (a.k.a., intruder, environment) may choose among many possible transitions in such a way so as to defeat security. This form of non-determinism is called *demonic* non-determinism.

Non-deterministic models have a well understood theory: algorithms for reachability, for model-checking against general modal formulae, for deciding bisimilarity are now part of the practitioner's knowledge.

Other models of computation are *probabilistic*: at any state of the computation, the choice of a next state obeys a given probability distribution. Particular kinds of such models are the *Markov chains* [5] and their variants, e.g., the labelled Markov processes [17], the *Markov decision processes* [44, 21] and the probabilistic programs of [34]. They are well-studied, too, and provide convenient mathematical tools: e.g., ergodic Markov chains have recurrent sets, i.e., there are methods that solve repeated reachability (à la Büchi). Notions of bisimulations (a.k.a., lumping [33]) for such systems have been investigated too [36, 13, 16].

Mixing both non-deterministic and probabilistic transitions has attracted a lot of interest in recent years. In the framework of transition systems, some of the most well-known approaches are the *Concurrent Labelled Markov Chains* and the *Probabilistic Automata*. See [57] for a comprehensive overview. The semantics foundations for these systems is a very active area of research, we mention in particular the approaches based on bisimulation [52, 43, 3, 15] and on metrics [19]. Model-checking such systems is a rather new activity as well (e.g., Marta Kwiatkowska's PRiSM tool [35]).

Security. We observe that modern security protocols must rely both on probabilistic choice and on demonic non-determinism. As a simple (and early) example, consider Chaum's dining cryptographers protocol [7]. In Chaum's colorful way of describing it, the problem is as follows. A group of at least three cryptographers $C_1, C_2, \ldots, C_n$ are sitting at a table in a restaurant. The waiter announces that their dinner has been paid for. Of course cryptographers must be extremely security-aware (some might think this is a politically correct way of saying "paranoid"). One of the cryptographers may have paid for the dinner, without wishing to reveal that he did, or it may be that the NSA has paid. The cryptographers want to decide which of the alternative is true. However, none of them is willing to reveal whether he indeed paid for the meal or not. The property that one cannot trace the payer is a form of *anonymity*. One of Chaum's solutions is as follows. Fix a random bit b_0, and give it to C_1. If C_1 has paid for the meal, then C_1 transmits $b_1 = \neg b_0$ to C_2, otherwise $b_1 = b_0$. When C_2 gets b_1, it transmits $b_2 = \neg b_1$ to C_3 if C_2 paid for the meal, $b_2 = b_1$ otherwise, and so on for $C_3, \ldots, C_n$. Eventually

C_n outputs a bit b_n. If $b_n = \neg b_0$, then one of the cryptographers indeed paid for dinner, and there is no way of telling who. Otherwise, somebody outside their group did, perhaps the NSA indeed.

Observe that Chaum's protocol critically depends on b_0 being random and uniformly distributed. Otherwise, C_i will be able to guess whether it is more likely that the payer is a C_j with $j < i$, or a C_k with $i < k$.

It should be clear that Chaum's protocol achieves its goal only if the bit b_i is transmitted to C_i in a secret and authenticated manner. This can be implemented by using electronic signatures and encryption algorithms. In this case, all actions of an intruder (external, or one of the cryptographers themselves) tending to forge messages and redirect communications—as in the Dolev-Yao model—can be thought as demonically non-deterministic: any sequence of intruder actions leading to an information leak with non-negligible probability counts as a successful attack.

On the importance of bisimulation. We have just talked about "information leak". The point is that, given a process $P(M)$ depending on the value of some piece of data M, an attacker will get some information about M if and only if there is some observable difference between $P(M)$ and $P(M_0)$, where M_0 is some other value. This reduces the *secrecy* of M to that of *observational equivalence* of $P(M)$ with $P(M_0)$. In the non-probabilistic setting, this is well-known, and is the cornerstone of the spi-calculus approach to secrecy [2]. Authentication can also be defined similarly, and notions of bisimulation exist that imply, and even are equivalent with observational equivalence [4].

Let us go beyond Chaum's dining cryptographers. Our first point is that an increasing number of current security properties rest on suitable notions of bisimulations. Notably, *anonymity*, *untraceability* (as in Chaum's protocol), *privacy*, *unlinkability*, and other *opaqueness*, a.k.a. *information hiding* properties are now required in various fields [29]. We refer to the latter paper for details, and we shall be content to say that applications include health-care applications (where illnesses and patient names are both public, so secrecy is not a required property, however the relation between patients and illnesses should remain unknown—except possibly for nation-wide statistics for use by physicians—this is untraceability), or electronic voting (where any voter should be able to verify the tally count without being able to deduce who voted for whom—this is again untraceability—, and it should be impossible for voters to give a convincing proof of whom they voted for, in order to prevent buying votes—this is another property called receipt-freeness or coercion-resistance) for example.

Previous approaches. While the examples in the previous paragraph have been studied in a non-probabilistic setting, such a setting is not satisfactory. We have already argued that protocols such as Chaum's dining cryptographers required adding some probabilistic reasoning. Examples abound: oblivious transfer, one-out-of-two oblivious transfer, bit commitment, coin flipping over the telephone are just a few examples of protocols that rely heavily on mixing probabilistic transitions with demonic non-determinism (as used by attackers). Note that oblivious transfer and bit commitment are used in zero-knowledge authentication systems [25]. In general, see [26, Chapter 11] for examples of such protocols, including verifiable secret sharing, electronic vote protocols, and digital cash protocols.

We have argued that we needed to use, and therefore needed to define first, frameworks for information hiding, based on suitable notions of bisimulation, and mixing (demonic) non-determinism with probabilities.

We have seen that most existing formal frameworks are either totally non-deterministic, or purely probabilistic. Traditional approaches to combine non-determinism and probability consist in eliminating one in favor of the other.

- Either probabilities are completely ignored, and random choices are replaced by non-deterministic choices. This is the approach used in most existing security models, e.g., in the spi-calculus. While this is adequate for simple deterministic protocols, we have seen that this was inadequate already for the simplest stochastic protocol, like Chaum's dining cryptographers.

- The other approach is to consider the non-deterministic dimension as some form of random choice with unknown probability. Typically, the kind of analysis based on these approaches is based on first resolving non-determinism and then calculating the probabilities. In the result of the analysis, the original non-determinism is represented by the two extreme cases, angelic (where probabilities of success are maximized) and demonic (where they are minimized). Estimating probabilities associated to non-deterministic choices is done e.g., by maximizing the entropy of some observable random variables. The latter is rather empirical, and is not guaranteed to account for the worst situation.

We wish to make an important and subtle point. In the ProNoBiS proposal, we stress mixing non-determinism and probabilities. One may think that Markov decision processes, labelled Markov processes, and Larsen-Skou processes [36] already include probabilities (clearly) and non-determinism (in the form of the choice of an action to execute next). We see those systems as deterministic: once the action is chosen, there is only one possible probability distribution along which to draw the next state. The fact that the weak logic $\mathcal{L}_0$ (see Section 4) already characterizes bisimulation equivalence in this setting testifies of this. It is true that one can always encode non-deterministic choices by giving action names to each alternative. This encoding preserves, say, reachability, but definitely not bisimulation. In contrast, we wish to express transitions where a given action gives rise to several possible probability distributions from which to draw the next state. Alternatively, this is equivalent to only allowing partial observations on actions taken.

Goal of the project. The ProNoBiS proposal emerged from discussions between the participants, who realized that they were independently starting to attack problems revolving around the combination of non-determinism and probability.

Accordingly, we envision at least two ways in which this problem may be tackled: using the so-called *theory of evidence* (see Section 1), and using the so-called *convex game theory* (see Section 2). Let us note that none of these theories are new, and that they share historical roots. There are several theories of evidence, the first being introduced by statisticians Dempster and Shafer in the 1960s [14, 53]. Convex game theory originates in economics, in the 1960s as well [55, 50]; the notion of games in this setting naturally goes back to Von Neumann and Morgenstern [62], but is in fact Choquet's notion of capacity [8], which generalizes that of measure. It should be noted that the totally monotone games encountered in economics are exactly Dempster and Shafer's belief functions [24, comment after Theorem 3.4]. This shows that there is, or at least used to be, some convergence between the two approaches, and we plan to compare them, see Section 3.

Our first step in this project is to realize that these are adequate models for mixing non-determinism and probabilities. Next, bisimulations have never been studied in these frameworks, to our knowledge, and we plan to find suitable definitions of the notion, and study algorithms to decide bisimulations. We shall develop later what needs to be done in these settings, in a computer science perspective.

In order to express security protocols which use randomization, we plan to develop a probabilistic process-calculus (see Section 5).

Bisimulation is one concern, as are applications to security, but we argue that expected benefits go beyond the domain of security. Applications to model-checking are discussed in Section 4, and applications to programming languages semantics are explored in Section 6.

1 WP 1: The Theory of Evidence

Participants: C. Palamidessi (leader), K. Chatzikokolakis, J. Goubault-Larrecq, V. Danos.

Assume that we have a number of hypotheses, mutually exclusive and exhaustive, and that each of these hypotheses leads to a probability distribution over a set of observations. The "evidence" is essentially a way to assign a "weight" to each of these hypotheses, on the basis of the observation. There are various definitions of evidence in literature (see [31] for a survey), some of which make sense only if there is a probability distribution also on the hypotheses. We are particularly interested in the case in which we do not assume any a-priori knowledge about the hypotheses, i.e. they are chosen non-deterministically.

We plan to consider the definition proposed by Shafer [54], later used also by Walley [63], and by Halpern and Fagin [28]. The idea is the following: Let $\mathcal{H}$ be the set of hypotheses, and $\mathcal{O}$ the set of observations. For each $h \in \mathcal{H}$, let μ_h be the probability distribution over $\mathcal{O}$ associated to h. Then, given $o \in \mathcal{O}$, the *evidence*, or *weight* of h is defined as:

$$w_o(h) = \frac{\mu_h(o)}{\sum_{h' \in \mathcal{H}} \mu_{h'}(o)}.$$

It is worth noting that, although conceptually different, w_o behaves as a probability distribution over H. Namely, $w_o(h)$ is a number between 0 and 1, and the sum of $w_o(h)$ over H is 1. This property is convenient from a technical point of view.

As an example, consider again Chaum's dining cryptographers protocol described in the introduction. Assume this time that the random bit generator is biased, and that it generates 0 and 1 with probabilities $2/3$ and $1/3$ respectively. For simplicity, let n

(the number of cryptographers) be 4. Assume that C_2 has not paid for the dinner, and that when he receives the bit from C_1, it observes that it has value 1. Later, C_2 comes to know (via the protocol) that one of the other cryptographers has paid. What is for C_2 the evidence $w_1(C_1)$ that C_1 is the payer? According to the above definition, it is

$$w_1(C_1) = \frac{\mu_{C_1}(1)}{\sum_{C \in \{C_1, C_3, C_4\}} \mu_C(1)} = \frac{2/3}{2/3 + 1/3 + 1/3} = 1/2$$

where $\mu_C(1)$ represents the probability that C_2 receives from C_1 the bit 1 when C is the payer.

A similar calculation shows that $w_1(C_i) = 1/4$ for $i = 3, 4$.

The theory of evidence seems a natural framework for reasoning about information-hiding properties and protocols, particularly in the case of anonymity. It is surprising that it has not been applied yet to this field.

We are aiming at establishing a framework for anonymity and information-hiding based on the theory of evidence. One of the first goals will be to revisit and define formally various informal anonymity notions that have been considered in literature. For instance, Reiter and Rubin's notion of "probable innocence" [45] would correspond to the fact that for the adversary the evidence that an user is "the culprit" is less than the evidence that he is not, under every observable.

2 WP 2: Capacities, Games, Belief Functions

Participants: J. Goubault-Larrecq (leader), C. Palamidessi, V. Danos.

A simple idea to try and reproduce as much as we can from Markov chain theory (i.e., without nondeterministic choice) is to model concurrent Markov chains by encoding both nondeterministic and probabilistic choice as a choice following some relaxed notion of random choice, where probabilities are replaced by some weaker notion. Let us call this relaxed notion "preprobability" for the time being.

We plan to use some forms of *capacities* (functions ν mapping sets to reals, only constrained by the fact that the preprobability of landing into the empty set is zero, i.e., $\nu(\emptyset) = 0$), and in particular *games* (which, in the economic literature, are monotone capacities, i.e., if $A \subseteq B$ then $\nu(A) \leq \nu(B)$).

Let us explain the basic intuition. Let A be a fixed set of states, and define informally the preprobability $P(q)(E)$ that, starting from state q, we land inside the set E, as 1 if and only if $A \subseteq E$, 0 otherwise. Clearly, $P(q)$ is *not* a probability—unless A has cardinal 1. However we claim that $P(q)$ encodes exactly the (demonic) nondeterministic choice of one successor q' among A. Imagine we play the role of P (process), and the adversary C (context) can choose any state $q' \in \mu(q) = A$. If A contains some state q' outside E, C can always choose to go to this state q', and ruin our hope of landing inside E. We see that the preprobability $P(q)(E)$ is the minimal probability that we land inside E, when C varies its choices. When $E \setminus A \neq \emptyset$, we have just seen that the adversary C can always force this probability to 0. On the other hand, if $A \subseteq E$, C must choose a state from A, hence in E, so $P(q)(E) = 1$.

This preprobability $P(q)$ is known as the *unanimity game* u_A on A [24], and is a generalization of Dirac masses (the Dirac mass at q' is $u_{\{q'\}}$). Unanimity games are not only monotone capacities, they are also *totally monotone*, in the sense that a generalized form of the inclusion-exclusion principle holds [24, Section 2, item (6)], and of course measures are also totally monotone games, a.k.a., *belief functions* in the sense of Dempster and Shafer [14, 53]. So belief functions appear to be reasonable grounds for founding the desired generalization of Markov chains to (labelled) concurrent Markov chains.

Adapting the classical notions of bisimulations, or of lumping, for Markov chains to this encoding of labelled concurrent Markov chains should be straightforward. Indeed, encode labelled concurrent Markov chains (with actions taken from some set L) as families $\mu_\ell : X \to B_1(X)$, $\ell \in L$, of transition functions mapping states in X to the preprobability distributions on possible successor states (as elements of the space $B_1(X)$ of normalized belief functions on X). When $\mu_\ell(q)$ is a probability, then this is a probabilistic transition; and when this is the case for every $q \in X$, we get an ordinary labelled Markov process. When $\mu_\ell(q)$ is a unanimity game u_A, then $\mu_\ell(q)$ encodes non-deterministic choice among states in A.

Now, in the purely probabilistic case where $\mu_\ell(q)$ is always a probability measure, a *lumping* [5] is any equivalence relation on X such that, for every $\ell \in L$, whenever $q_1 \equiv q_2$, for every equivalence class C closed under $\equiv$, $\mu_\ell(q_1)(C) = \mu_\ell(q_2)(C)$. This is easily seen to be equivalent to the definition given by Larsen and Skou of probabilistic bisimulation [36]. And we observe that the definition makes absolutely no use of the fact that $\mu_\ell(q_1)$ or $\mu_\ell(q_2)$ is a probability. For that matter, general capacities would be acceptable.

It is not however clear at the moment of this writing that this notion would enjoy all expected properties of a bisimulation, e.g., that bisimilar systems are observationally equivalent, or that some Hennessy-Milner-like logic could characterize bisimilar systems. It is our plan to study such properties.

To do so, we shall need to study the foundations of the theory of belief functions and related notions, as a first step. The survey paper [24] gives a good overview of most of the theorems that we need, in the case of finite-state systems.

However, we shall also be interested in infinite-state systems. This is particularly clear if we consider applications to security protocols, where spaces of messages exchanged between principals are described as terms from some infinite term algebra (possibly modulo an equational theory), as in the Dolev-Yao model. More generally, it would be beneficial to be able to rest future studies on a corpus of general theorems that would apply to transition systems with continuous state spaces X (as in [16] for labelled Markov processes)—i.e., to general enough classes of topological spaces. Resting this foundational study on topology rather than measure theory seems to provide for opportunities of having unified view. E.g., terms can be seen as the finite part of a cpo of terms with undefined parts, and cpo theory is a subdomain of topology. In this study, we would extend the theory of continuous valuations initiated by Saheb-Djahromi [48, 49] and Jones [30] to continuous games, and continuous belief functions.

3 WP 3: Relation Between the Two Approaches

Participants: C. Palamidessi (leader), K. Chatzikokolakis, J. Goubault-Larrecq, V. Danos.

We have said in the introduction that there have been some crossovers between the various theories of evidence proposed in the past and research in economics on capacities and games. The most prominent being the fact that Dempster and Shafer's works [14, 53] are cited in both contexts.

At the moment, we do not understand fully the relations between the two approaches, and one aim of this proposal is to understand precisely each other's point of view. For example, belief functions qua totally monotone games (Section 2) were proposed by Shafer in a book entitled "A Mathematical Theory of Evidence", suggesting a link to the approach of Section 1. However, belief functions do not involve computing evidences or weights as in Section 1, where the notion of evidence originates from another paper of Shafer's [54]. A theory of belief based on the latter has been developed by Halpern and Fagin [28].

In economics, often a more useful notion than belief functions are convex games. Every belief function is a convex game, and every convex game is a game. (Inclusions are strict.) We shall argue in Section 6 that convex games should be studied in their own right from a computer science perspective.

These are questions that we should answer in the framework of this proposal.

4 WP 4: Model-Checking and Temporal Logics

Participants: M. Kwiatkowska (leader), C. Palamidessi, J. Goubault-Larrecq, R. Segala, V. Danos, L. Fribourg, C. Picaronny, S. Pinot, Ph. Schnoebelen, N. Bertrand, A. Parma, A. Turrini, P. Wu.

In non-probabilistic approaches, quotienting the transition system under study by the largest bisimulation is a well-known way of reducing the size of the system under study. Since this preserves all properties expressible in standard modal logics, quotienting helps model-checking. In other words, let I be a Kripke model, and φ a modal formula. If $\equiv$ is the largest bisimulation, then $I \models \varphi$ if and only if $I/\equiv \models \varphi$, and the latter is in general easier to compute. In the domain of Markov chains (where probabilities are present but not non-determinism), bisimulations can be used for the same purpose; this is well-known, and this is called lumping [33].

More generally, we plan to develop new model-checking techniques, not limited to reachability properties.

This will first involve defining adequate logics. One avenue which is particularly elegant from the mathematical point of view is to try and find logics that characterize bisimulation equivalence. One early example is Hennessy-Milner logic, in the purely non-deterministic setting. Another example is the logic $\mathcal{L}_0$ [16], which characterizes bisimulation in the case of purely probabilistic systems—namely, labelled Markov processes. $\mathcal{L}_0$ is a surprisingly weak logic, since it does not contain negation, disjunction, or infinitary constructions. (It has finite conjunctions and formulas of the form $\langle \ell \rangle_q F$, meaning that we can follow a transition labelled ℓ and arrive at some state satisfying F with probability at least q.) We plan to look for similar logics characterizing bisimulation equivalence in the framework of labelled concurrent Markov chains. This, in turn, can be used to define and compute bisimulation *distances* so that any two processes are at distance 0 if and only if they are bisimilar; the idea was implemented in [18, 60] for labelled Markov processes and later refined by Ferns and Panangaden [22]. The relation between distances and least modal depths of formulae satisfied by one system and not the other was exploited in [11, 10, 12], again for labelled Markov processes.

In practice, being able to compute distances allows one to evaluate whether two systems are reasonably close; in security, whether anonymity is reasonably preserved. Insisting on exact bisimulation is in general absurd, since two systems with the same transition P, one with probability 0.5, the other with probability 0.49999999 for example, cannot be bisimilar", write ": bisimilarity is not *robust*. But, in some sense, these two systems should be close. Another avenue to define a notion of closedness is through topology, and in particular Scott approximants, and we plan to explore both avenues and understand the relationships between the two.

Even logics that are complete for bisimulation equivalence can be weak; see the $\mathcal{L}_0$ example. It will therefore be interesting to define other, more expressive logics. For example, one may add fixpoint operators [9]. In each case, we plan to explore model-checking algorithms. Such algorithms may benefit from lumping to reduce the state space, or from approximations, as alluded to above.

5 WP 5: A Formalism to Express Security Protocols

Participants: R. Segala (leader), C. Palamidessi, J. Goubault-Larrecq, V. Danos, K. Chatzikokolakis, A. Parma, A. Turrini, R. Beauxis.

One of the goals of our proposal is to develop a formalism to represent security protocols and systems. We aim at a language in the style of a process calculus, because it will allow to benefit from the rich suite of results and tools developed in the field of Concurrency Theory. In particular, we are interested in modeling probabilistic protocols, so the calculus will have to incorporate mechanisms to express probabilistic choice.

Most existing process calculi for security lack of probabilistic constructs. One exception is the probabilistic version of the spi-calculus [37]. However, the language in [37] is obtained by replacing the parallel operator of the spi-calculus [2] with a probabilistic parallel operator. There is no choice operator. Consequently, the language is purely probabilistic.

In our proposal, on the contrary, we want to consider both a probabilistic choice, which can be regarded as a random choice made by the process, and a standard parallel operator, which is controlled by the scheduler. The scheduler can be nondeterministic or probabilistic, depending on the protocol and on the properties that we want to model.

The advantage of our approach, to our opinion, is that the scheduler should be considered an entity separated from the process, and possibly cooperating (or controlled) by the adversary. We believe that it is necessary, in general, to ensure that the properties of the protocols are verified under as little assumptions as possible concerning the scheduler.

Our starting point will be the probabilistic asynchronous π-calculus developed in [42]. However, that language contains input-guarded probabilistic choice and that is perhaps more complicated than what we need for our purposes: a purely internal probabilistic choice suffices to represent all the randomized security protocols we know of. On the other hand, the language in [42], just like the π-calculus, is very basic. For expressing protocols it will be convenient to introduce value passing, primitive functions, and data types like it is done in the *applied π-calculus* [1]

We plan to to write an interpreter for this language and to establish its foundations using the model of probabilistic automata, and the semantic studies developed in the other WPs of this proposal.

6 WP 6: Semantics of Programming Languages

Participants: V. Danos (leader), J. Goubault-Larrecq, R. Harmer

Another expected benefit of the approaches of Section 1 and Section 2 is to provide new semantics for higher-order languages with both non-determinism and probabilistic choice. This is an old problem in the semantics of programming languages, which boils down to the fact that one cannot commute non-deterministic choices with probabilistic choices.

The question of finding a good way of combining non-determinism and probabilistic choice in the semantics of programming languages is an old problem. We believe that the theory of convex or concave games provides a good solution to this problem. Preliminary studies indicate that totally monotone games, i.e., belief functions, as described in Section 2, and their duals, plausibility functions, provide a poor basis for such a task. Convex games, on the other hand, enjoy a form of the Riesz Representation Theorem, which states that the map sending a game ν on X to its integration functional $f \mapsto \int_{x \in X} f(x) d\nu$ is an isomorphism between convex games and so-called colinear lower previsions (we adapt some terminology from Maaß [39]): this is Schmeidler's Theorem [51]. And it seems that lower previsions (not necessarily colinear) enjoy all required properties to form a monad, with strong connections to convex games. Also, expressing semantics with this monad would essentially allow one to state the semantics of expressions as weakest pre-expectations, i.e., as reward functions.

We propose to explore this and compare this to previous approaches.

Let us give a slightly more detailed picture of research in this domain. We have mentioned the word "monad", and indeed we propose to examine the semantics of programming languages under the lens of Moggi's computational lambda-calculus [41], where side-effects (here, non-deterministic and probabilistic choices) are described by a strong monad on a cartesian-closed category. In our case, this means we would study a lambda-calculus with both non-deterministic and probabilistic choice. Not only do monads provide a particularly elegant approach to the semantics of such languages, but we could then use the results of [27] to derive so-called logical relations for this language. One application is that any two programs that are related by some logical relation are automatically observationally equivalent. This would yield, in particular, a framework for security in the presence of non-determinism, probabilistic choice, and higher-order computations.

Also, it was claimed in [27] and some further papers that logical relations (or extensions: Kripke logical relations, sconing constructions, fibrations) provide an extension of the notion of bisimulation to higher-order computations. Note that labelled concurrent Markov chains would just be computational λ-terms of type $Q \to L \to \boldsymbol{T}Q$, where Q is the type of states, L is the type of labels, and $\boldsymbol{T}$ is the monad type constructor. We would therefore like to compare this approach to the construction of bisimulations of Section 2.

A relatively general construction of a monad combining the two effects of a monad for non-determinism and a monad for probabilistic choice, in a given category, is still subject to discussion. However, Lüth proposed in his PhD thesis [38] that the right construction was the *coproduct* of the two monads, which exists under mild conditions [32]. However, in general, the coproduct of two monads is a relatively inscrutable object. In special cases, a simpler description is available, e.g., for two *ideal* monads [23]. Notably, the coproduct of the *non-blocking* non-determinism monad (where the set of possible choices of C is never empty) with the probability monad falls into this case. As can be expected, the resulting monad is formed of all sequences of choices, alternating between probabilistic and non-deterministic choices [23, Example 4.3].

Varacca also proposed a monad that integrates both non-deterministic and probabilistic choice [61]. Ghani and Uustalu [23] note that the coproduct monad is very close to the notion of synchronization trees that Varacca uses.

In programming language semantics, it is often necessary to replace sets by complete partial orders (cpos). The fact that we can generalize game and belief function theory to continuous spaces, as we aim to do (see Section 2), will definitely be useful here.

Using cpos, it is well-known that three notions of non-determinism emerge: demonic (as in concurrent Markov chains, angelic (where C tries to help you get the highest reward), and chaotic. We shall be mostly interested in demonic non-determinism; Varacca's monad, and therefore also Ghani and Uustalu's, mixes probabilities with *angelic* non-determinism. We suspect that some variants of the theory of continuous games should apply to angelic non-determinism as well, but this won't be central to this proposal.

Mislove was probably the first to solve the general problem of mixing non-determinism with probabilities [40]. His solution is elegant, and considers free cpos over given equational theories. This covers all forms of non-determinism, but maybe lacks some concrete feel. Tix also solved the problem in her PhD thesis [58, 59], while covering again all forms of non-determinism. Her solution only works on specific cpos called *d-cones*, and involves considering convex lower powercones, convex upper powercones, and biconvex powercones respectively, which are particular sets of probabilities. We plan to compare these approaches to ours. It seems that the convex lower powercone should correspond rather precisely to the space of all convex games, as suggested by Shapley's Theorem [55, 56] and Rosenmuller's Theorem [46, 47] in the finite case.

7 Profile of the Principal Investigators

Vincent Danos, University of Paris 7, France Vincent Danos is Directeur de Recherches at the CNRS in the PPS lab which he contributed to create in 2000. His main body of work since 2003—inasmuch as it relates to the present proposal—concerns: probabilistic models in a general measure-theoretic framework with particular attention to compression techniques, and the certification of transactional mechanisms. He has also more recently been touching on the topic of distributed quantum protocols which also include a form of probabilistic evolution adn security concerns.

He has a long-standing collobaration with: Prakash Panangaden (Prof., McGill) and Josée Desharnais (Assistant Prof., Laval) who both extended the study of probabilistic models to general state spaces and developed over the last decade powerful approximation techniques for such models; Jean Krivine (Doc, INRIA) in the study of distributed consensus in reversible process algebras; and Elham Kashefi (post-doc, IQC and Oxford), Ellie d'Hondt (post-doc, Vrije U. Brussels) in the study of distributed quantum protocols.

He has published about 50 papers in Theoretical Computer Science and Mathematical Logic, has been serving as a PC member for more than 20 international conferences, is an editor of LMCS (Logical Methods in Computer Science) and TCSB (Transactions on Computational Systems Biology), was the chair of CMSB04 (Computational methods in Systems Biology), and is a member of the IFIP Working group (WG1.8) on Concurrency Theory.

Jean Goubault, ENS Cachan, France Jean Goubault is a former student of Ecole Polytechnique where he was ranked first at the entry exam and again first at the closing exam. He is currently holding of position of Professor at the LSV lab in ENS Cachan. He has a long track record in various topics in Theoretical Computer Science covering subjects such as proof-theory, automated proof-search, tree automata, model-checking, domain-theoretic and category-theoretic semantics of higher-order languages, programming language design, and computer algebra software. He has written a book in Proof-theory and Automated deduction, has been member of many PC in international conferences and is leading various projects related to security and formal methods.

Marta Kwiatkowska, University of Birmingham, UK Marta Kwiatkowska is Professor of Computer Science in the School of Computer Science at the University of Birmingham. Prior to that she was Assistant Professor at the Institute of Computer Science, at the Jagiellonian University in Krakow and Lecturer at the Department of Mathematics and Computer Science at the University of Leicester.

Marta Kwiatkowska, has an internationally leading reputation in probabilistic verification. She works on modelling languages and process calculi, verification algorithms and software tools, and applications of verification technology to real world case studies, ranging from Internet and mobile ad hoc network protocols, embedded systems and ubiquitous computing devices, to biological organisms. Kwiatkowska is Principal Investigator of research projects totalling £1M and participated in the APPSEM EU network.

A major achievement of Kwiatkowska's group has been the development of the foundations for probabilistic verification and implementation of the internationally leading probabilistic model checker PRISM (`www.cs.bham.ac.uk/~dxp/prism/`). PRISM has been used to model and analyse real-world protocols such as IEEE 1394 FireWire root contention, IEEE 802.11 WLAN, Bluetooth device discovery, IPv4 Zeroconf link-local addressing, probabilistic anonymity and contract signing, and reliability of nanotechnology circuit designs. PRISM is a focus of international interest, with users and contributors worldwide, for example at Stanford, CMU, LSV Cachan, Edinburgh, Rome, Virginia Tech, KTH, UNSW and Monash.

Jointly with colleagues Mark Ryan and Georgios Theodoropolous, Kwiatkowska's research is also directed at applications of model checking to industrially relevant problems, specifically the feature interaction problem and verification of asynchronous hardware (collaboration with Manchester).

Catuscia Palamidessi, INRIA, FRANCE Catuscia Palamidessi is Director of Research at INRIA Futurs, where she leads the team Comète. She got her PhD at the University of Pisa in 1988. She worked as Full Professor at the University of Genova, Italy (1994-1997) and at the Pennsylvania State University, USA (1998-2002).

Catuscia Palamidessi's research interests include Concurrency, Mobility, and Distributed Systems. Her past achievements include the proof of expressiveness gaps between various concurrent calculi, and the development of a probabilistic version of the asynchronous π-calculus. Her current research is in mobile calculi, probability, and the use of probabilistic concepts in Concurrency and in Security.

Catuscia Palamidessi has been appointed Program Committee Chair of 5 conferences, including CONCUR 2000 and ICALP 2005, she has been invited speaker at other 5 conferences, including CONCUR'99 and PPDP 2003, and she has served as PC member in more than 50 conferences. She is in the Editorial board of the CUP journal TPLP (Theory and Practice of Logic Programming) and of Elsevier's ENTCS (Electronic Notes in Theoretical Computer Science). She is in the Steering Committee of EATCS (the European Association of Theoretical Computer Science) and PPDP (Principles and Practice of Declarative Programming).

Roberto Segala, Verona University, ITALY Roberto Segala received a Laurea in Scienze dell'Informazione in 1991 as a student of the Scuola Normale Superiore in Pisa. In 1992 he received a Master in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology, and in 1995 he received a PhD in Computer Science from the same institution. From 1995 to 2001 he was Assistant Professor (ricercatore) at the University of Bologna; from 2001 to 2005 he was Associate Professor at the University of Verona, and since 2005 he is Professor of Computer Science at the University of Verona.

The main research interests of Roberto Segala concern the study of models of concurrency extended with paradigms like real-time, hybrid behaviour, and stochastic behaviour, and their use in the analysis and verification of concurrent and distributed systems.

Some of the most important results obtained by Roberto Segala are the study of the connections between the theory of testing and the verification methods based on I/O automata, the study of receptiveness for liveness properties in the real-time settings, the definition of (nondeterministic) Probabilistic Automata, now known as the non alternating model for stochastic nondeterministic systems, the definition of weak and branching bisimulation relations in the context of stochastic systems, the compositional analysis of randomized distributed algorithms. Other results obtained from international collaborations are the extension of timed automata with discrete probabilistic transitions and the study of model checking procedures on the extended model, and the definition of hybrid I/O automata for the compositional analysis of hybrid systems.

Roberto Segala is a member of the Steering Committee of QEST (Quantitative Evaluation of SysTems) and is a member of the editorial board of the International Journal of Hybrid Systems. He performs regular editorial activity as a member of the Program Committee of several international conferences and he was invited to several conferences and international schools where he presented his research results.

References

[1] Martín Abadi and Cédric Fournet. Mobile values, new names, and secure communication. In *Proc. 28th ACM Symposium on Principles of Programming Languages (POPL'01)*, pages 104–115, 2001.

[2] Martín Abadi and Andrew D. Gordon. A calculus for cryptographic protocols: The spi calculus. *Information and Computation*, 148(1):1–70, 1999.

[3] Emanuele Bandini and Roberto Segala. Axiomatizations for probabilistic bisimulation. In *Proceedings of the 28th International Colloquium on Automata, Languages and Programming*, volume 2076 of *Lecture Notes in Computer Science*, pages 370–381. Springer, 2001.

[4] Johannes Borgström and Uwe Nestmann. On bisimulations for the spi calculus. In Hélène Kirchner and Christophe Ringeissen, editors, *Proc. 9th Int. Conf. Algebraic Methodology and Software Technology (AMAST'2002), Saint-Gilles-les-Bains, La Réunion, France, sep. 2002*, pages 287–303. Springer Verlag LNCS 2422, 2002.

[5] Pierre Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer-Verlag, New York, 1999.

[6] Giuseppe Castagna, Jan Vitek, and Franco Zappa Nardelli. The Seal calculus. *Information and Computation*, 201:1–54, 2005.

[7] David L. Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1(1):65–75, 1988.

[8] Gustave Choquet. Théorie des capacités. *Annales de l'Institut Fourier*, 5(4):131–295, 1953.

[9] Vincent Danos and Josée Desharnais. A fixpoint logic for labeled Markov Processes. In Zoltan Esik and Igor Walukiewicz, editors, *Proceedings of the international Workshop Fixed Points in Computer Science, FICS'03*, Warsaw, 2003.

[10] Vincent Danos and Josée Desharnais. Labeled Markov Processes: Stronger and faster approximations. In *Proceedings of the 18^{th} Symposium on Logic in Computer Science, LICS'03*, Ottawa, 2003. IEEE Computer Society Press.

[11] Vincent Danos, Josée Desharnais, and Prakash Panangaden. Conditional expectations and the approximation of labeled Markov Processes. In *Proceedings of CONCUR'03, Marseille, France*, volume 2761 of *Lecture Notes in Computer Science*, pages 477–491. Springer-Verlag, 2003.

[12] Vincent Danos, Josée Desharnais, and Prakash Panangaden. Labeled Markov Processes: Stronger and faster approximations. *Electronic Notes in Theoretical Computer Science*, 87:157–203, September 2004.

[13] Erik P. de Vink and Jan J. M. M. Rutten. Bisimulation for probabilistic transition systems: A coalgebraic approach. *Theoretical Computer Science*, 221(1–2):271–293, 1999.

[14] Arthur P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

[15] Yuxin Deng and Catuscia Palamidessi. Axiomatizations for probabilistic finite-state behaviors. In *Proceedings of the 8th International Conference on Foundations of Software Science and Computation Structures*, volume 3441 of *Lecture Notes in Computer Science*, pages 110–124. Springer, 2005.

[16] Josée Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for labelled Markov processes. *Information and Computation*, 2002. To appear.

[17] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximating labelled Markov processes. *Information and Computation*, 184(1):160–200, 2003.

[18] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004.

[19] Josee Desharnais, Radha Jagadeesan, Vineet Gupta, and Prakash Panangaden. The metric analogue of weak bisimulation for probabilistic processes. In *Proceedings of the 17th Annual IEEE Symposium on Logic in Computer Science*, pages 413–422. IEEE Computer Society, 2002.

[20] Danny Dolev and Andrew C. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, IT-29(2):198–208, 1983.

[21] Eugene A. Feinberg and Adam Schwartz, editors. *Handbook of Markov Decision Processes, Methods and Applications*. Kluwer, 2002. 565 pages.

[22] Norm Ferns, Prakash Panangaden, and Doina Precup. Metrics for markov decision processes with infinite state spaces. Internal report, available at `http://rl.cs.mcgill.ca/~prakash/uai05.pdf`.

[23] Neil Ghani and Tarmo Uustalu. Coproducts of ideal monads. *Theoretical Informatics and Applications*, 38(4):321–342, 2004. Extended abstract in Z. Ésik, I. Walukiewicz, ed., Proc. of Int. Workshop on Fixed Points in Computer Science, FICS'03 (Warsaw, Apr. 2003), pp. 32-36. Warsaw Univ., 2003.

[24] Itzhak Gilboa and David Schmeidler. Additive representation of non-additive measures and the Choquet integral. Discussion Papers 985, Northwestern University, Center for Mathematical Studies in Economics and Management Science, 1992.

[25] Oded Goldreich, Silvio Micali, and Avi Wigderson. Proofs that yield nothing but their validity and a methodology of cryptographic protocol design. In *Proc. 27th IEEE Symposium on Foundations of Computer Science (FOCS'86)*, pages 174–187. IEEE Computer Society Press, 1986.

[26] Shafi Goldwasser and Mihir Bellare. *Lecture Notes on Cryptography*. Available at `http://www.cs.ucsd.edu/users/mihir/papers/gb.html`, 2001.

[27] Jean Goubault-Larrecq, Slawomir Lasota, and David Nowak. Logical relations for monadic types. In *Proceedings of the 16th International Workshop on Computer Science Logic (CSL'02)*. Springer-Verlag LNCS 2471, 2002.

[28] Joseph Y. Halpern and Ronald Fagin. Two views of belief: Belief as generalized probability and belief as evidence. *Artificial Intelligence*, 54(3):275–317, April 1992.

[29] Dominic Hughes and Vitaly Shmatikov. Information hiding, anonymity and privacy: A modular approach. *Journal of Computer Security*, 2003. To appear.

[30] Claire Jones. *Probabilistic Non-Determinism*. PhD thesis, University of Edinburgh, 1990. Technical Report ECS-LFCS-90-105.

[31] Henry E. Kyburg Jr. Recent work in inductive logic. In *Recent Work in Phylosophy*, pages 87–150. Rowman & Allanheld, 1983.

[32] G. Max Kelly. A unified treatment of transfinite constructions for free algebras, free monoids, colimits, associated sheaves and so on. *Bulletin of the Australian Mathematical Society*, 22:1–83, 1980.

[33] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. Springer-Verlag, New York, 1976.

[34] Dexter Kozen. Semantics of probabilistic programs. *Journal of Computer and Systems Sciences*, 22:328–350, 1981.

[35] Marta Kwiatkowska, G. Norman, and D. Parker. PRISM: Probabilistic symbolic model checker. In P. Kemper, editor, *Proc. Tools Session of Aachen 2001 International Multiconference on Measurement, Modelling and Evaluation of Computer-Communication Systems*, pages 7–12, September 2001. Available as Technical Report 760/2001, University of Dortmund.

[36] Kim G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94:1–28, 1991.

[37] P. Lincoln, J. Mitchell, M. Mitchell, and A. Scedrov. A probabilistic poly-time framework for protocol analysis. In *Proceedings of the 5th ACM Conference on Computer and Communications Security*, pages 112–121, San Francisco, California, November 1998. ACM Press.

[38] Christoph Lüth. *Categorical Term Rewriting: Monads and Modularity*. PhD thesis, University of Edinburgh, 1997.

[39] Sebastian Maaß. Coherent lower previsions as exact functionals and their (sigma-)core. In *Proceedings of the 2nd Intl. Symp. Imprecise Probabilities and their Applications (ISIPTA'01)*, pages 230–236, Maastricht, the Netherlands, 2001.

[40] Michael Mislove. Nondeterminism and probabilistic choice: Obeying the law. In *Proc. 11th Conf. Concurrency Theory (CONCUR'00)*, pages 350–364. Springer Verlag LNCS 1877, 2000.

[41] Eugenio Moggi. Notions of computation and monads. *Information and Computation*, 93:55–92, 1991.

[42] Catuscia Palamidessi and Oltea M. Herescu. A randomized encoding of the π-calculus with mixed choice. In *Proceedings of the 2nd IFIP International Conference on Theoretical Computer Science*, pages 537–549, 2002.

[43] Anna Philippou, Insup Lee, and Oleg Sokolsky. Weak bisimulation for probabilistic systems. In *Proceedings of the 11th International Conference on Concurrency Theory*, volume 1877 of *Lecture Notes in Computer Science*, pages 334–349. Springer-Verlag, 2000.

[44] Martin L. Puterman. *Markov Decision Processes–Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, NY, 1994.

[45] Michael K. Reiter and Aviel D. Rubin. Crowds: anonymity for Web transactions. *ACM Transactions on Information and System Security*, 1(1):66–92, 1998.

[46] J. Rosenmuller. On core and value. *Methods of Operations Research*, 9:84–104, 1971.

[47] J. Rosenmuller. Some properties of convex set functions, part II. *Methods of Operations Research*, 17:287–307, 1972.

[48] Nasser Saheb-Djahromi. Probabilistic LCF. In *Mathematical Foundations of Computer Science*, volume 64. Springer-Verlag, 1978.

[49] Nasser Saheb-Djahromi. CPO's of measures for non-determinism. *Theoretical Computer Science*, 12(1):19–37, 1980.

[50] Herbert E. Scarf. The core of an N person game. *Econometrica*, 35(1), 1967. Cowles Foundation Paper 277.

[51] David Schmeidler. Integral representation without additivity. *Proceedings of the American Mathematical Society*, 97:255–261, 1986.

[52] Roberto Segala and Nancy Lynch. Probabilistic simulations for probabilistic processes. *Nordic Journal of Computing*, 2(2):250–273, 1995. An extended abstract appeared in *Proceedings of CONCUR '94*, LNCS 836: 481-496.

[53] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, USA, 1976.

[54] Glenn Shafer. Belief functions and parametric models (with commentary). *Journal of the Royal Statistical Society, Series B*, 44:322–352, 1982.

[55] Lloyd S. Shapley. Notes on n-person games VII: Cores of convex games. The Rand Corporation R. M., 1965.

[56] Lloyd S. Shapley. Cores of convex games. *International Journal of Game Theory*, 1:12–26, 1971.

[57] A. Sokolova and E. P. de Vink. Probabilistic automata: system types, parallel composition and comparison. In *Validation of Stochastic Systems: A Guide to Current Research*, volume 2925 of *Lecture Notes in Computer Science*, pages 1–43. Springer, 2004.

[58] Regina Tix. *Continuous D-Cones: Convexity and Powerdomain Constructions*. PhD thesis, Technische Universität Darmstadt, 1999.

[59] Regina Tix, Klaus Keimel, and Gordon Plotkin. Semantic domains for combining probability and non-determinism. *Electronic Notes in Theoretical Computer Science*, 129:1–104, 2005.

[60] Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 331(1):115–142, 2005.

[61] Daniele Varacca. The powerdomain of indexed valuations. In *Proc. 17th Annual IEEE Symp. on Logic in Computer Science (LICS'02)*, pages 299–308, Copenhagen, July 2002. IEEE Computer Society Press.

[62] John Von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ, USA, 1944.

[63] Peter Walley. Belief function representations of statistical evidence. *Annals of Statistics*, 18(4):1439–1465, 1987.